ATARI®

A Warner Communications Company

Indice

1 Introduzione

Il programma **STATISTICA I** della Atari è stato progettato per elaborare i dati inseriti in un archivio e fornisce dati statistici in generale.
In un'unica elaborazione è possibile ottenere la media, la moda, la mediana, la deviazione standard, la varianza, la dissimetria, la curtosi e lo scarto quadratico medio di un insieme di valori.
Le funzioni di editing, integrali al programma, permettono la correzione o la sostituzione di dati errati. I dati possono essere vantaggiosamente memorizzati su cassetta o dischetto e poi caricati direttamente nella memoria del sistema.

Negli ultimi anni la statistica ha subito una notevole evoluzione passando da una semplice rappresentazione grafica dei dati, normalmente chiamata "statistica descrittiva", ad un metodo più scientifico e preciso che si basa su un attento esame dei dati disponibili in modo da ricavare da essi previsioni e decisioni. Questo metodo più evoluto, spesso chiamato "statistica induttiva", è basato sulla manipolazione dei dati attraverso formule matematiche e sul calcolo delle probabilità. Le iterazioni matematiche, che nella maggior parte dei casi sono di banale semplicità, spesso si riducono ad esecuzioni noiose di operazioni ripetitive. Ed è appunto in questa sfera di attività che l'elettronica dei calcolatori trova uno dei suoi migliori impieghi.
Il calcolatore, usato come un veloce, preciso ed affidabile "number cruncher" o meglio, un masticatore di numeri, assume un ruolo importantissimo. Quasi indispensabile.

Naturalmente, non tutte le decisioni necessitano di o si basano su elaborazioni intimamente collegate alla statistica; alcune possono essere immediate ed intuitive. Molte conclusioni sono ovvie, anche se un esame affrettato dei dati a volte può trarre in inganno anche l'occhio più addestrato.
Non sarà certo la prima volta che un'assunzione un po' affrettata risulta poi errata quando si viene a conoscenza di tutti i fattori statistici che concorrono alla definizione del fenomeno.
Si consideri, per esempio, il caso di due studenti che frequentano scuole diverse e che svolgono entrambi un compito in classe di matematica.
Il primo, Gianni, ottiene il voto 78 mentre il suo amico Roberto, che studia in un'altra scuola, ottiene 86.
Senza ombra di dubbio, chiunque asserirebbe che Roberto è chiaramente più bravo in matematica. Sicuramente a prima vista questo è vero.
Vogliamo analizzare il caso un po' più nei dettagli?

Supponiamo dunque di sapere che il voto medio per la classe di Gianni è 60, con una deviazione standard di 12, mentre il voto medio per la classe di Roberto è 74 con una deviazione standard di 10. E' possibile quindi asserire che la votazione di Gianni ha una deviazione standard di (78-60)/12=1,5 rispetto al voto medio della propria classe, mentre quella di Roberto ha la deviazione standard (86-74)/10=1,2 rispetto al voto medio della propria classe. Risulterebbe da questa prima analisi che Gianni, rispetto alla propria classe, è uno studente migliore di Roberto!
Come mai la prima conclusione che sembrava a prima vista così scontata risulta invece non così corretta?

Probabilmente le difficoltà dell'esame erano diverse, come potevano essere anche diversi i contenuti. Il fatto determinante rimarrà però che, disponendo di informazioni addizionali e ricorrendo ad un metodo statistico, si possono raggiungere conclusioni più significative che a volte non coincidono con ipotesi formulate empiricamente.
L'adozione della metodologia statistica dipende interamente dal tipo di dato o di informazione che si ha a disposizione. E' indispensabile quindi provvedere ad una raccolta accurata dei dati se si vuole ricavare una qualunque conclusione significativa o, per lo meno, indicativa.
Per esempio, volendo determinare le probabilità che ha un cavallo di vincere una particolare corsa, facendo riferimento ai suoi risultati precedenti, non bisognerà prendere in considerazione indiscriminatamente tutte le corse a cui il cavallo ha partecipato, ma solo quelle che accomunano il maggior numero di caratteristiche.

Un dato può essere un valore numerico qualunque ma, generalmente, rappresenta un valore di "quanto grande" (ampiezza) o "quante volte" (frequenza). Una problematica che si presenta molto frequentemente è quella di ottenere dati che esprimono frequenza ricavata da dati che sono collegati al concetto di ampiezza. Ciò può essere facilmente risolto dividendo equamente l'intero intervallo dei punti o dei valori in gruppi, chiamati intervalli di classe, e contando i punti che cadono in ogni intervallo.
Questo numero rappresenta la frequenza del punto.

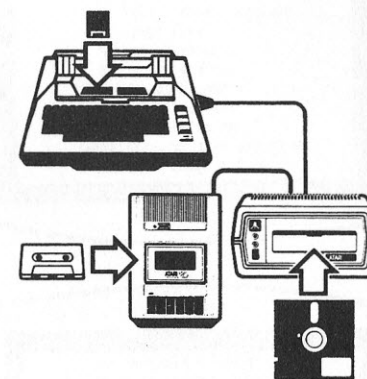
Per quanto riguarda il programma STATISTICA I dell'Atari si assume che i dati abbiano una casualità normale, spesso chiamata "curva a campana". La distribuzione normale è largamente usata nella teoria delle probabilità. Le equazioni usate nei calcoli sono spiegate nell'Appendice A.

Il programma STATISTICA I nei calcoli statistici utilizza equazioni standard di uso comune.
Se un altro calcolatore o un altro metodo statistico dovesse produrre risultati diversi per un'elaborazione statistica, sarà opportuno confrontare le equazioni.

2 Caricamento del Programma

Il programma viene fornito in due versioni, una per disco e l'altra per cassetta.

- Il programma per cassetta occupa una sola cassetta.
- Il programma per disco occupa un solo dischetto.

Per le istruzioni di caricamento, tanto del disco quanto della cassetta, si rimanda ai manuali operativi delle due macchine ed alla relativa Appendice.

Note

3 Il Funzionamento

Il programma inizia la sua esecuzione visualizzando il simbolo della Atari, il nome del programma e le informazioni inerenti ai diritti d'autore.
Dopo una breve permanenza a video, appare il menu del programma, composto da sette opzioni.

STATISTICA I - VERSIONE SU CASSETTA
ARCHIVIO - NONOME
A: INPUT NUOVO ARCHIVIO
B: EDIT DATI
C: SAVE ARCHIVIO
D: LOAD ARCHIVIO
E: ELABORA I DATI
F: STAMPA I DATI
G: RITORNA AL BASIC
INSERISCI L'OPZIONE DESIDERATA
(LETTERA DALLA A ALLA G):
DIRITTI RISERVATI (C) 1980 ATARI

STATISTICA I - VERSIONE SU DISCO
ARCHIVIO - NONOME
A: INPUT NUOVO ARCHIVIO
B: EDIT DATI
C: SAVE ARCHIVIO
D: LOAD ARCHIVIO
E: ELABORA I DATI
F: STAMPA I DATI
G: RITORNA AL BASIC
INSERISCI L'OPZIONE DESIDERATA
(LETTERA DALLA A ALLA G)
DIRITTI RISERVATI (C) 1980 ATARI

La scelta di un'opzione avviene introducendo da tastiera la lettera corrispondente e premendo **RETURN**.
L'identificazione generica attribuita ad un archivio sarà "NONOME".

Opzioni del Menu

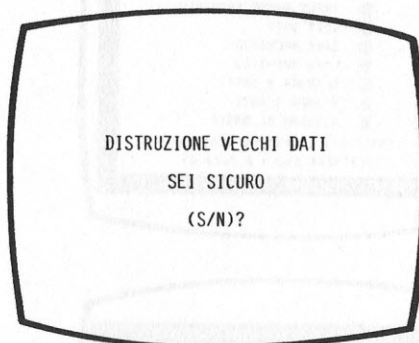
Le opzioni disponibili vengono presentate operativamente nello stesso ordine in cui appaiono nel menu.

Opzione A -Input Nuovo Archivio

La scelta di questa opzione permette di inserire un nuovo insieme di punti o dati da analizzare.
Se i dati vengono inseriti per la prima volta dopo il caricamento del programma, la visualizzazione verrà modificata per facilitare l'inserimento dei dati.
Se il programma è stato eseguito ed in memoria sono presenti i vecchi dati, vengono emessi quattro segnali acustici e a video viene richiesta un'operazione di verifica al fine di ridurre la possibilità di cancellazione accidentale di dati utili.

DISTRUZIONE VECCHI DATI

SEI SICURO

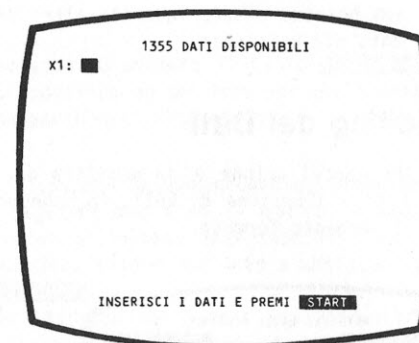
(S/N)?

Se a tale richiesta si risponde S seguito da RETURN, i dati vecchi vengono cancellati e a video viene presentato il prospetto relativo all'INSERIMENTO; se si risponde N e poi RETURN, i dati vecchi vengono conservati e si ritorna al menu.
Il prospetto di INSERIMENTO ha la forma illustrata nella figura che segue.

NOTA

1355 è il numero dei dati disponibili; l'esatto valore dipende naturalmente dalla quantità di memoria RAM di cui dispone il sistema.
Il programma necessita 16K bytes per essere eseguito.

1355 DATI DISPONIBILI

X1: ■

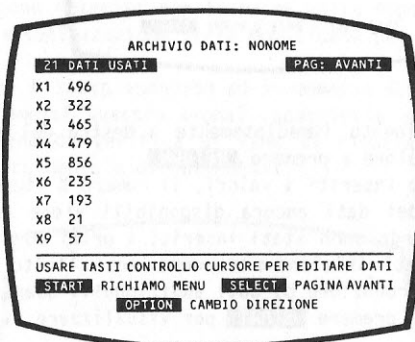
INSERISCI I DATI E PREMI START

Il cursore è posizionato immediatamente a destra del primo dato X1. Inserire il primo valore e premere RETURN.
Man mano che vengono inseriti i valori, il numero X viene incrementato di 1 e il numero dei dati ancora disponibili viene automaticamente diminuito di 1. Quando sono stati inseriti i primi 20 valori gli identificatori X dei dati iniziano a scorrere verso l'alto dello schermo per permettere l'introduzione di dati addizionali. Quando tutti i dati sono stati inseriti, premere START per visualizzare il menu.

L'inserimento di ciascun dato deve essere seguito da un RETURN. Se durante l'introduzione viene accidentalmente digitato un carattere non numerico, il sistema emette un segnale acustico e nella parte alta dello schermo visualizza il messaggio **DATO NON VALIDO**.
Ogni numero valido per il **BASIC** è valido anche per il programma di **STATISTICA I**. Comunque, dato che i numeri vengono elevati al quadrato o al cubo durante i calcoli, il valore dei dati **NON** dovrebbe superare 5E+24 o essere inferiore a -5E+24.
Se viene usato tutto lo spazio disponibile, qualunque tentativo di inserire altri numeri provoca un segnale acustico e apparirà a video la seguente visualizzazione:

In queste condizioni non possono essere inseriti altri dati senza prima cancellare qualche dato già esistente.

Opzione B - Editing dei Dati

Permette l'aggiunta, la cancellazione e la modifica di dati già esistenti. Quando viene scelta l'opzione di EDIT, lo schermo visualizza i dati esistenti con il seguente formato.

ARCHIVIO DATI: NONOME
21 DATI USATI   PAG: AVANTI
X1 496
X2 322
X3 465
X4 479
X5 856
X6 235
X7 193
X8 21
X9 57
USARE TASTI CONTROLLO CURSORE PER EDITARE DATI
START RICHIAMO MENU   SELECT PAGINA AVANTI
OPTION CAMBIO DIREZIONE

Il menu B, nella fase di visualizzazione, fornisce le seguenti funzionalità:

- La totalizzazione a video dei dati usati con l'aggiornamento progressivo in funzione di cancellazioni o nuovi inserimenti.
- La visualizzazione di 9 dati alla volta presentati in sequenza.
- L'utilizzo dell'Opzione PAG: AVANTI o PAG: INDIETRO.
    - PAG: AVANTI visualizza i successivi 9 dati in ordine cresente. Premere **SELECT** per avanzare.
    - PAG: INDIETRO visualizza i precedenti 9 dati in ordine decrescente. Premere **OPTION** per cambiare direzione.
- Non è permesso l'uso dell'Opzione PAGE XXXX oltre l'inizio o la fine dei dati.

Azionando un qualsiasi tasto diverso da **OPTION**, **SELECT** o **START** cambia la funzione a video da visualizzazione a modo EDITING DEI DATI per aggiornare gli stessi. Due modi sono disponibili:

1) - Posizionare il cursore tramite i tasti di controllo sopra i dati da variare. Inserire i nuovi valori e premere **RETURN**.

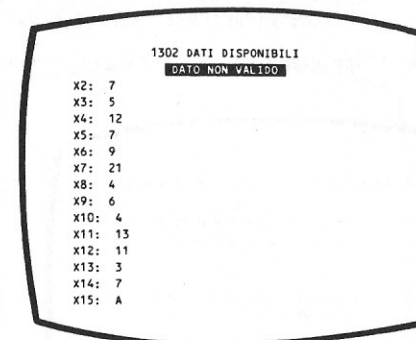
oppure:

2) - Digitare il numero X di identificazione del dato da cambiare. Inserire il nuovo valore preceduto da uno spazio e poi **RETURN**.

NOTA

E' necessario premere il tasto **RETURN** dopo una qualsiasi modifica ad un dato per abilitarne l'accettazione da parte del calcolatore.

Il modo 2 illustrato precedentemente, permette anche l'aggiunta di nuovi dati. Digitare una X ed il numero di identificazione immediatamente successivo all'ultimo inserimento, inserire uno spazio bianco, il valore che deve essere aggiunto e abilitare l'entrata del dato con il comando **RETURN**.
Se non viene lasciato uno spazio tra l'identificatore X e il valore o se nel valore del dato viene inserito un carattere non numerico (diverso dal simbolo esponenziale E), nella parte superiore dello schermo viene prodotto un messaggio di errore e viene generato un segnale acustico. All'insorgere di tali condizioni il dato non viene modificato e sarà necessario ripetere la sequenza delle operazioni in modo corretto.

1302 DATI DISPONIBILI
DATO NON VALIDO
X2: 7
X3: 5
X4: 12
X5: 7
X6: 9
X7: 21
X8: 4
X9: 6
X10: 4
X11: 13
X12: 11
X13: 3
X14: 7
X15: A

Un dato può essere cancellato digitando X e il suo numero di riferimento, ad esempio X5, e premendo **RETURN**. Questa operazione causa la cancellazione del punto X5 e lo scorrimento verso l'alto dei dati successivi all'X5 cancellato. Il dato che era X6 assume il nuovo valore di X5, X7 diventa X6, e così via. L'ultimo identificatore X della catena viene cancellato e la totalizzazione dei dati utilizzati viene ridotta di 1.
Se un dato viene cancellato per errore esso può essere aggiunto al set di dati, ma solo in coda. I nuovi dati non possono essere inseriti tra i dati esistenti, ma possono solo essere aggiunti in coda alla catena dei dati. Comunque, ciò non influenza affatto i calcoli statistici. Se invece, per una qualsiasi ragione, la progressività è importante, un modo di inserire un dato in una posizione particolare è quello di modificare via via ogni punto a partire dal punto in errore.

Oppure si può ritornare al menu, scegliere con **A** l'Opzione INPUT NUOVO ARCHIVIO, cancellare i dati inseriti e inserire l'intero set di dati nell'ordine desiderato.

Opzione C – Salvataggio Archivio su Cassetta o Dischetto

1 - SALVATAGGIO SU CASSETTA

E' sempre consigliabile memorizzare i propri archivi su una cassetta diversa da quella che contiene il programma STATISTICA I. Le ragioni per cui si dovrebbe ricorrere al salvataggio di un archivio sono:

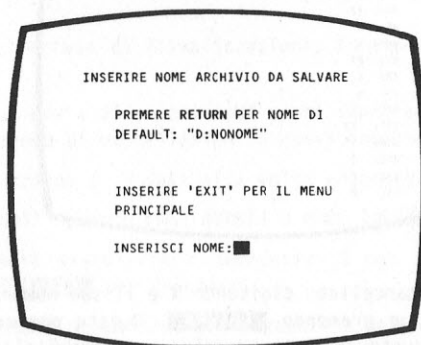
- L'elevata dimensione del file,
- La ripetitività delle operazioni, con eventuali ritorni al BASIC tra elaborazioni,
- Previsioni di future aggiunte di dati,
- Creazione di una libreria di riferimento per utilizzi futuri.

Quando viene scelta l'Opzione C di salvataggio, viene visualizzato:

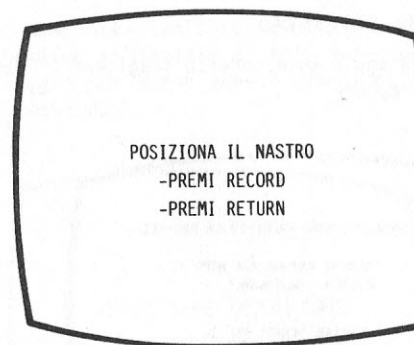
```
INSERIRE NOME ARCHIVIO DA SALVARE

PREMERE RETURN PER NOME DI
DEFAULT: "D:NONOME"

INSERIRE 'EXIT' PER IL MENU
PRINCIPALE

INSERISCI NOME:█
```

Se si vuole assegnare un nome all'archivio creato:

- Tornare al Menu Principale, digitando EXIT, seguito da **RETURN**,
- Inserire un nome non superiore a 10 caratteri, dato che ogni carattere inserito oltre al decimo, viene ignorato. Questo nome verrà adottato come identificazione definitiva per quella catena di dati e usato per futuri richiami,
- Premere **RETURN**, per convalidare il nuovo nome.


```
POSIZIONA IL NASTRO
-PREMI RECORD
-PREMI RETURN
```

Poichè il tasto RECORD sul Registratore di Programmi non è autobloccante quando usato da solo, bisogna azionare contemporaneamente i tasti RECORD e PLAY per salvare l'archivio su nastro. Due segnali acustici ricordano poi di sbloccare i due tasti.
Dopo aver premuto **RETURN** per iniziare il salvataggio, controllare, attraverso la finestrella del registratore, che il nastro sia in moto. Al termine della registrazione su nastro, appare a video:

```
ARCHIVIO
D:NOME
SALVATO
```

Questo messaggio rimane a video per pochi secondi, quindi riappare sullo schemo il Menu Principale.

NOTA

Assicurarsi che il nastro sia posizionato correttamente per evitare l'accidentale sovrascrittura su dati già esistenti e che i tasti RECORD e PLAY siano rilasciati prima di dare il **RETURN**.

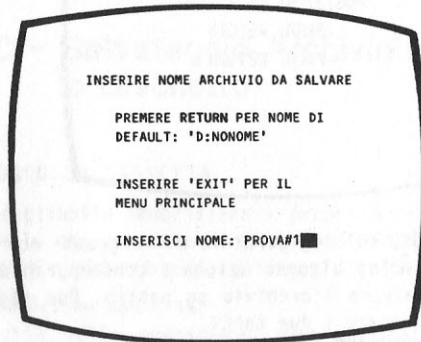
2 - SALVATAGGIO SU DISCHETTO

Per eseguire un salvataggio su dischetto scegliere l'Opzione C seguita da **RETURN**. A video appare:

```
INSERIRE NOME ARCHIVIO DA SALVARE

    PREMERE RETURN PER NOME DI
    DEFAULT: 'D:NONOME'

    INSERIRE 'EXIT' PER IL
    MENU PRINCIPALE

    INSERISCI NOME: PROVA#1█
```

A questo punto l'utente può scegliere di salvare i suoi dati su di un nuovo dischetto, oppure sul dischetto che contiene il programma STATISTICA I. Quando il nuovo nome è stato introdotto, appare a video il messaggio:

```
ARCHIVIO
D:PROVA#1
SALVATO
```

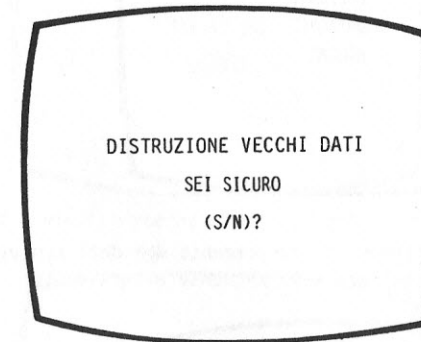
Questo messaggio rimane a video per pochi secondi, dopo di che il computer restituisce a video il Menu Principale.

Opzione D - Caricamento Archivio da Cassetta o Dischetto

1 - CARICAMENTO DA CASSETTA

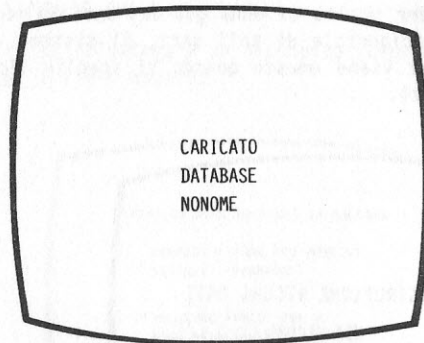
Se un archivio è stato precedentemente salvato su cassetta con il Registratore di Programmi, la scelta dell'Opzione D permette di ricaricarlo in memoria in modo da poterlo listare, modificare, elaborare.

Se quando l'opzione viene scelta ci sono già dei dati in memoria, per prevenire la perdita accidentale di tali dati, il sistema visualizza lo stesso messaggio che viene emesso quando si sceglie l'opzione di inserimento di nuovi dati.

```
DISTRUZIONE VECCHI DATI
SEI SICURO
(S/N)?
```

La risposta N provoca il ritorno al Menu Principale e il mantenimento dei dati attualmente presenti in memoria. La risposta S causa immediatamente l'emissione di un segnale acustico, seguito dalla visualizzazione delle seguenti tre righe:

```
POSIZIONA IL NASTRO
-PREMI PLAY
-PREMI RETURN
```

Posizionare il tasto all'inizio del file contenente l'archivio da caricare utilizzando il tasto ADVANCE sul registratore e il contagiri se si conosce la posizione iniziale dell'archivio sul nastro. Altrimenti occorre riavvolgere il nastro fino all'inizio con il tasto REWIND e iniziare il caricamento del primo archivio presente su nastro. Bisognerà caricare ogni file in successione finché non si incontra l'archivio ricercato. Una volta che il nastro è correttamente posizionato, l'operazione di caricamento viene attivata azionando **RETURN**. Se il caricamento è stato eseguito correttamente, appare a video:

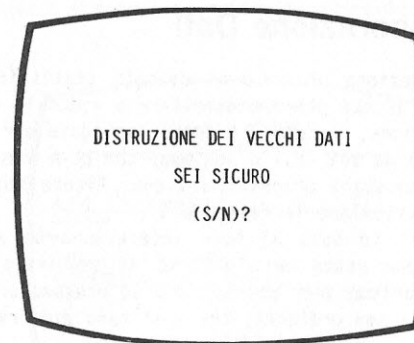
CARICATO
DATABASE
NONOME

Se per una qualsiasi ragione il caricamento dei dati non avviene in modo corretto, viene visualizzato il seguente avvertimento:

ERRORE NEL REGISTRATORE
A CASSETTA

In entrambi i casi viene visualizzato il Menu Principale per effettuare ulteriori scelte. Se il caricamento non è terminato con successo riposizionare attentamente il nastro e ripetere l'operazione.

2 - CARICAMENTO DA DISCHETTO

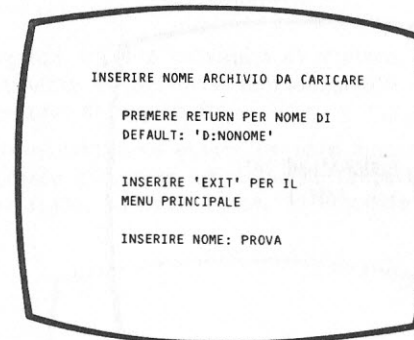
Il caricamento di un archivio da dischetto è molto simile al caricamento da cassetta, eccezione fatta che l'operazione di posizionamento del nastro è sostituita dalla digitazione del nome dell'archivio da caricare. L'operazione di caricamento viene attivata digitando D seguito da **RETURN**. Se ci sono già dei dati in memoria, a video si vedrà:

DISTRUZIONE DEI VECCHI DATI
SEI SICURO
(S/N)?

La risposta N causa il ritorno al menu, mentre S visualizza:

INSERIRE NOME ARCHIVIO DA CARICARE

PREMERE RETURN PER NOME DI
DEFAULT: 'D:NONOME'

INSERIRE 'EXIT' PER IL
MENU PRINCIPALE

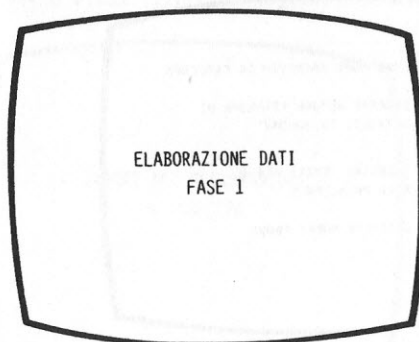
INSERIRE NOME: PROVA

Dopo aver inserito il nome di un archivio precedentemente salvato e aver azionato **RETURN**, a video si leggerà:

ARCHIVIO
D:PROVA
CARICATO

Questo messaggio rimane a video per pochi secondi, dopo di che viene visualizzato il Menu Principale.

Opzione E - Elaborazione Dati

La scelta di questa opzione dà avvio ai calcoli statistici sui dati presenti in memoria. L'utente potrà controllare a video il susseguirsi delle fasi dell'esecuzione, durante il tempo impiegato per i calcoli. Il processo è composto da sei fasi e man mano che esse vengono eseguite vengono variati i messaggi proposti a video. Alcune fasi sono più lunghe di altre, in particolare le fasi 4 e 5.
L'ordinamento dei punti in base al loro valore assorbe molto tempo elaborativo poiché viene usato un algoritmo di sequenzializzazione. E' stato scelto un algoritmo per ottimizzare le prestazioni del computer quando si hanno dati non ordinati, che è il caso più frequente.

```
ELABORAZIONE DATI
     FASE 1
```

Le sei fasi corrispondono approssimativamente ai seguenti passi:

1) Calcolo della media.
2) Calcolo dello scarto quadratico medio, deviazione standard, varianza e curtosi.
3) Preparazione dei vettori usati nell'operazione di ordinamento.
4) Ordinamento dei dati.
5) Ricerca delle mode.
6) Determinazione della mediana e calcolo dell'asimmetria.

Se il numero dei punti è elevato, il calcolatore impiega un po' di tempo per calcolare i valori statistici. Dopo che sono passati 9 minuti senza che sia stato azionato alcun tasto, il calcolatore inizia a ciclare su tutti i colori disponibili. Questo processo può essere arrestato premendo un tasto qualsiasi della tastiera.

Quando la Fase 6 è terminata, appare il seguente risultato:

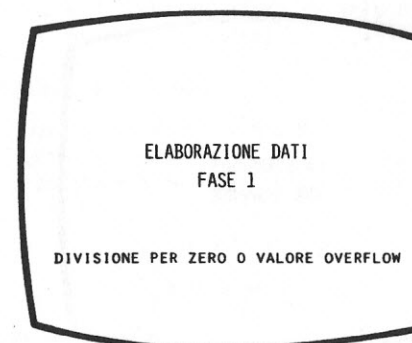
```
       ARCHIVIO - NONOME
     STATISTICA GENERALE
      23 DATI ELABORATI
MEDIA   =  8.69565217
MEDIANA =  7
S.Q.M.  =  10.6648549
DEV.SEV    = 6.31329604
VARIANZA   = 39.85770688
CURTOSI    = 8.55693344
ASIMMETRIA = 1.14595033
MODE:
 3 7

  PREMI OPTION PER STAMAPRE LA STATISTICA
      PREMI START PER CONTINUARE
```

Se una stampante Atari è collegata al sistema, la pressione del tasto OPTION permette di ottenere il risultato su carta. La stampa avrà lo stesso formato della videata.

Se non ci sono dati o se esiste un solo dato o il valore di un dato è troppo grande o troppo piccolo, nella parte inferiore del video viene visualizzato, su campo rosso, il seguente messaggio:

```
      ELABORAZIONE DATI
           FASE 1

DIVISIONE PER ZERO O VALORE OVERFLOW
```

Dopo una breve pausa, il programma ripropone il Menu Principale.

Opzione F – Stampa Dati

La scelta di questa opzione provoca la stampa dei dati presenti in memoria sulla stampante Atari. Durante la stampa appare a video il messaggio ELENCO DEI DATI. Il formato della stampa è il seguente.

ELENCO DELL'ARCHIVIO "NONOME"

X1 : 3	X11 : 13	X21 : 3
X2 : 7	X12 : 11	X22 : 36
X3 : 5	X13 : 3	X23 : 5
X4 : 12	X14 : 7	
X5 : 7	X15 : 8	
X6 : 9	X16 : 18	
X7 : 21	X17 : 10	
X8 : 4	X18 : 14	
X9 : 6	X19 : 2	
X10 : 4	X20 : 2	

- Se si vuole arrestare la stampa dopo che essa ha avuto inizio, è sufficiente premere il tasto **START**
- Se viene azionato il tasto **OPTION** e la stampante non è disponibile il sistema emette un segnale acustico per indicare l'errore
- Se la stampante non è connessa o disponibile, lo schermo visualizza:

LA STAMPANTE
NON RISPONDE

Dopo pochi secondi il Menu Principale viene riproposto a video.

Opzione G – Ritorno al BASIC

Quando viene scelta questa opzione del menu, si determina l'uscita immediata dal programma STATISTICA I ed il BASIC riprende il controllo del computer. Tutti i dati in memoria relativi al programma statistico vengono persi a meno che non siano stati preventivamente salvati su cassetta o dischetto.

4 Appendice A

I calcoli statistici del programma si basano sulle seguenti formule:

1 - MEDIA

X = (X1 + X2 + X3 +...+ Xn)/N

dove:

X1, X2 ... sono i dati e N è il numero totale dei dati.

2 - MEDIANA

Dopo aver ordinato i dati in ordine ascendente, dal più piccolo al più grande, la **MEDIANA** è il valore del dato centrale nel caso in cui il numero dei dati sia dispari. Quando invece il numero dei dati è pari, la **MEDIANA** è il valore di mezzo tra i due dati centrali del set di dati elaborati.

3 - MODA

La **MODA** è il valore o i valori che si presentano con frequenza maggiore. In un insieme di punti ci possono essere più mode o nessuna moda se tutti i valori sono presenti una sola volta.

4 - SCARTO QUADRATICO MEDIO

Lo Scarto Quadratico Medio o **S.Q.M.** è la radice quadrata della somma di tutti i termini X, elevati al quadrato, e poi divisa per il numero totale N dei termini X considerati:

$$R.M.S. = \sqrt{\frac{\sum_{i=1}^{N} (x_i)^2}{N}}$$

5 - DEVIAZIONE STANDARD

La **DEVIAZIONE STANDARD** è la radice quadrata della somma di tutte le differenze tra il generico dato x_i e il valore medio, elevato al quadrato, e divisa per il numero totale dei dati -1.

NOTA

Nei campioni di dati utilizzano N-1 al denominatore. I valori statistici vengono comunemente calcolati su un campione piuttosto che sull'intero insieme, cioè tutti i dati possibili.

$$s = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N-1}}$$

6 - VARIANZA

La **VARIANZA** è il quadrato della deviazione standard.

$$s^2 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N-1}$$

7 - CURTOSI

La **CURTOSI** è il valore di picco rispetto ad una distribuzione normale. La curva **NORMALE** viene generalmente considerata come una curva di picco uguale a 3. Quindi, una distribuzione con picco maggiore di 3 viene detta curva a **CURTOSI** maggiore del normale e viene detta leptocurtica.
Se il picco è minore di 3, la curva viene detta platicurtica. Una distribuzione esattamente normale viene detta mesocurtica ed ha curtosi uguale a 0.
Matematicamente la curtosi viene definita come una funzione di quarto grado della **MEDIA** ed è calcolata come somma delle differenze di ogni termine **X** e il valore medio, ognuna elevata alla quarta potenza, divisa prima per il numero totale dei termini **X**, cioè **N**, e poi per la potenza quarta della deviazione standard e sottraendo poi 3 dal risultato finale.

$$a_4 = \left[\frac{\sum_{i=1}^{N} (x_i - \bar{x})^4}{N \cdot s^4} \right] - 3$$

8 - ASIMMETRIA

L'**ASIMMETRIA** è la misura della non simmetria della distribuzione rispetto ad una distribuzione normale ed è calcolata come una funzione cubica della **MEDIA**. E' uguale alla somma del cubo delle differenze tra ogni termine x_i e la media, per tutti i termini x, e poi divisa per il numero dei termini x, cioè **N**. Questo risultato viene ulteriormente diviso per la deviazione standard ed elevato al cubo. Per una curva normale il suo valore è zero. Una asimmetria positiva indica che la maggior parte dei punti stanno a sinistra rispetto al valore medio. Un valore negativo indica che i punti sono raggruppati a destra della media.

$$a_3 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^3}{N \cdot s^3}$$

5 Appendice B

ESEMPIO 1

Viene fornito un esempio di utilizzo del programma **STATISTICA I** della Atari.

Si supponga di essere un piccolo commerciante e di avere, ad esempio, un negozio di attrezzature sportive. Si è notato che i clienti normalmente affluiscono al negozio in numero diverso durante il corso della giornata; spesso molti clienti affollano il negozio con un conseguente disservizio.
Si potrebbe ricorrere all'assunzione di un'altra persona per velocizzare il servizio durante le ore di punta, ma non esiste la certezza che il giro degli affari giustifichi questo onere addizionale.
Chiaramente prima di rendere questa decisione operativa sarà opportuno rispondere a domande che chiariscano la dinamica del negozio e le aspettative economico-gestionali del titolare.

L'orario di apertura del negozio è dalle 7 del mattino alle 11 di sera e una domanda perfettamente giustificabile potrebbe essere se economicamente vale la pena rimanere aperti così a lungo.
L'andamento degli affari è abbastanza positivo nella prima mattinata e nel tardo pomeriggio per tenere aperto il negozio durante queste ore?
Quanti clienti in media entrano nel negozio durante la giornata?
Quanto cambia questa media se il negozio rimane aperto per un periodo più breve durante il giorno?
Quale è il tempo medio per soddisfare ciascun cliente?
Quale è il ricavo medio di ogni vendita?
Quante tra le persone presenti nel negozio ad un determinato istante di tempo sono clienti che pagano?

Le domande sono molte ma, poiché tutte possono essere esaminate con criteri similari, solamente due verranno prese in considerazione ed esaminate. La lista che segue rappresenta il numero dei clienti presenti nel negozio, raggruppati per ciascuna ora, durante una giornata con apertura di 16 ore.
I valori statistici più interessanti sono la **MEDIA**, la **DEVIAZIONE STANDARD**, la **ASIMMETRIA** e la **CURTOSI**.

La **MEDIA** indica che 6 persone sono mediamente presenti nel negozio ogni ora, ma la **DEVIAZIONE STANDARD** di 4,6 è un valore troppo grande rispetto alla media. Ciò significa che esiste una grande variazione nel numero di clienti da un'ora all'altra. La **CURTOSI** di -3,1 indica

che il picco della curva è più piatto rispetto a quello di una curva normale e che la presenza dei clienti è piuttosto distribuita su tutto l'arco della giornata. L'**ASIMMETRIA** pari a 0,858 mostra che il picco della curva non è centrato; la maggior parte dei clienti entra nel negozio prima dell'ora centrale della giornata, ossia prima delle 3 pomeridiane.

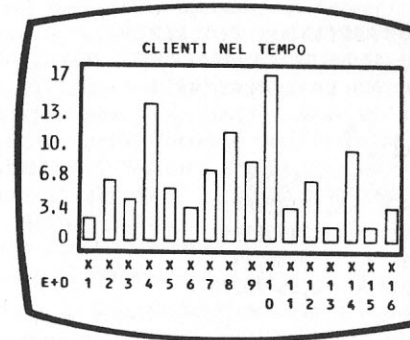
ELENCO DELL'ARCHIVIO "D: CLIENTI"

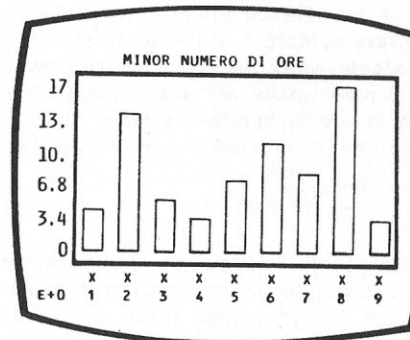
X1 : 2
X2 : 6
X3 : 4
X4 : 14
X5 : 5
X6 : 3
X7 : 7
X8 : 11
X9 : 8
X10 : 17
X11 : 3
X12 : 6
X13 : 1
X14 : 9
X15 : 1
X16 : 3

ARCHIVIO - D: CLIENTI
STATISTICA GENERALE
16 DATI ELABORATI

MEDIA = 6, 25
MEDIANA = 5,5
S.Q.M. = 7,68927825
DEV. STD = 4,6260134
VARIANZA = 21,39999997
CURTOSI = -0,30509181
ASIMMETRIA= 0,8579832608
MODE : 3

Il grafico rappresentato qui di seguito illustra il numero dei clienti presenti ogni ora nel negozio con apertura a tempo pieno dalle 7 di mattina alle 11 di sera.
I dati visualizzati sotto forma di istogramma forniscono una visualizzazione immediata dell'andamento della curva e dei valori di picco. I grafici presentati in questa Appendice sono stati generati dal programma **GRAFICA** della Atari.

CLIENTI NEL TEMPO
17
13.
10.
6.8
3.4
0
E+0 X1 X2 X3 X4 X5 X6 X7 X8 X9 X10 X11 X12 X13 X14 X15 X16

Si supponga ora di ridurre l'orario di esercizio aprendo il negozio alle 9 del mattino e chiudendolo alle 18 pomeridiane.

MINOR NUMERO DI ORE
17
13.
10.
6.8
3.4
0
E+0 X1 X2 X3 X4 X5 X6 X7 X8 X9

Dopo aver inserito nel sistema ed elaborato i dati corrispondenti alle ore scelte si ottiene:

ELENCO DELL'ARCHIVIO "D:CLIENTI"

X1 : 4
X2 : 14
X3 : 5
X4 : 3
X5 : 7
X6 : 11
X7 : 8
X8 : 17
X9 : 3

ARCHIVIO - D:CLIENTI
STATISTICA GENERALE
9 DATI ELABORATI

MEDIA = 8
MEDIANA = 7
S.Q.M. = 9,29755045
DEV. STD = 5,02493781
VARIANZA = 25,24999999
CURTOSI = -1,33985774
ASIMMETRIA = 0,5517037917
MODE : 3

Con questi nuovi presupposti, la **MEDIA** dei clienti per ora è passata da 6 a 8. La **DEVIAZIONE STANDARD** è salita a 5,0 ed è più significativa rispetto al valore 8 di quanto non lo fosse 4,6 per 6.
La **CURTOSI** è diminuita a -1,34 ed indica che la curva dei clienti per ora è ancora più bassa.
Si noti che l'**ASIMMETRIA** è passata da 0,858 a 0,552 e ciò indica che il picco della curva è in effetti più vicino al suo centro, mentre prima, il picco si trovava molto più a sinistra della **MEDIA**.
Naturalmente, questo accade anche perchè il punto centrale della giornata è passato dalle 3 pomeridiane all'una e mezza. Questi risultati indicano che riducendo le ore di apertura durante la giornata, il numero dei clienti per ora che si prevede viene incrementato del 29,13%:

$$\frac{8-6,25}{6} = 29,17\%$$

Chiaramente, quanto esposto non rappresenta una situazione che può essere generalizzata; lo scopo della discussione era puramente quello di illustrare l'utilizzo del programma e non quello di fornire una soluzione completa.

ESEMPIO 2

Si supponga di pianificare un viaggio per andare a pesca nelle Valli di Comacchio. L'intenzione è quella di accamparsi con una tenda nel terreno di qualche contadino compiacente. Si supponga pure di voler andare durante il mese di aprile in coincidenza con l'apertura della stagione della pesca.
L'unica grande preoccupazione è abbastanza comprensibile.
"Come sarà il tempo?"
Uno dei fattori che ci piacerebbe conoscere è la temperatura media e l'escursione media.

Le temperature qui di seguito elencate si riferiscono al mese di aprile nella zona di Comacchio e sono rappresentative delle medie giornaliere, espresse per ogni giorno del mese.
Si è particolarmente interessati al valore medio ed alla deviazione standard. Dopo aver inserito tutti i dati ed averli elaborati usando

l'**OPZIONE** E del menu del programma, si troverà che la temperatura media per aprile è 18,5°C (65,3°F) con **DEVIAZIONE STANDARD** di 12,79°C (8,98°F). Questo indica che un mese di aprile tipico è moderatamente freddo, con oscillazioni di temperatura di modesta entità.
Ciò non significa però, che tutti i mesi di aprile siano moderati, e proprio questo aprile potrebbe rappresentare l'eccezione; d'altronde, i dati statistici indicano che ci sono buone probabilità di tempo buono. In base a tali dati, si possono prevedere temperature che si aggirano sui 18,4°C (65,0°F).
Un altro dubbio potrebbe essere:
"Pioverà durante il campeggio?"

Utilizzando il programma **STATISTICA I** con i dati relativi alle precipitazioni in un mese di aprile tipico, si potrà avere una buona previsione sul comportamento metereologico del mese.

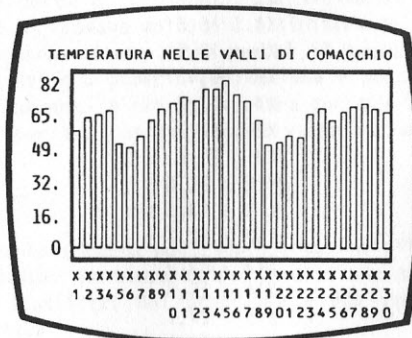
ELENCO DELL'ARCHIVIO 'D:TEMPERCOM'

X1 : 14,4
X2 : 17,7
X3 : 18,8
X4 : 16,6
X5 : 10,5
X6 : 9,4
X7 : 12,7
X8 : 17,2
X9 : 20,5
X10 : 21,6
X11 : 23,3
X12 : 22,7
X13 : 25,5
X14 : 25,8
X15 : 27,7
X16 : 23,8
X17 : 22,7
X18 : 17,2
X19 : 10,0
X20 : 10,7
X21 : 12,2
X22 : 12,7
X23 : 18,8
X24 : 20,0
X25 : 17,2
X26 : 19,4
X27 : 21,1
X28 : 21,6
X29 : 20,0
X30 : 19,3

ARCHIVIO - D:TEMPERCOM
STATISTICA GENERALE
30 DATI ELABORATI

MEDIA = 65,32
MEDIANA = 66,9
S.Q.M. = 65,91387815
DEV. STD. = 8,97910985
VARIANZA = 80,62441369
CURTOSI = = -0,91550317
ASIMMETRIA = -0,2628555067
MODE : 63

Rappresentazione grafica delle temperature durante il mese di aprile nelle Valli di Comacchio.

TEMPERATURA NELLE VALLI DI COMACCHIO
82
65.
49.
32.
16.
0

6 Manuali di Riferimento

1. Young, Hugh D. : Trattato Statistico e Dati Sperimentali
   McGraw-Hill Book Company, Inc., New York, 1962

2. Spiegel, Murray R. : Statistiche (Note Generalizzate di Schaum)
   McGraw-Hill Book Company, Inc., New York, 1961

3. Nie, Norman H., e altri : Pacchetto Statistico per Scienze Sociali
   McGraw-Hill Book Company, Inc., New York, 1975

4. Acton, Forman S., : Metodi Numerici di Lavoro
   Harper & Row, Publishers, New York, 1960

5. Carnahan, Brice, e altri : Metodi Numerici Applicati
   John Wiley & Sons, Inc., New York, 1979

6. Freund, John E., : Statistica - Propedeutico
   Prentice-Hall, Inc., New Jersey, 1970